

This is a working document prepared by the Energy Information Administration (EIA) in order to solicit advice and comment on statistical matter from the American Statistical Association Committee on Energy Statistics. This topic will be discussed at EIA's spring 2006, meeting with the Committee to be held April 6 and 7, 2006.

Using Performance Statistics to Design the 2006 Manufacturing Energy Consumption Survey

Robert K. Adler (robert.adler@eia.doe.gov ph:202-586-1134) and Thomas G. Lorenz(thomas.lorenz@eia.doe.gov ph:202-586-3442)

Objectives

This paper examines some of the performance statistics available from the conduct of the 2002 Manufacturing Energy Consumption Survey (MECS). First, we look at a sample of data status flags to see if they lead to useful results about evaluating the effectiveness of the questionnaire. Second, we compare the status flag data for respondents who completed an Excel version of the questionnaires to the data for those respondents who completed a written version. Together, those two sections point to the desirability of an electronic version of the 2006 MECS. Finally, we examine differential response rate to suggest other changes that may be made to improve data quality in the MECS.

Background

The Manufacturing Energy Consumption Survey (MECS) is the U.S. Government's official survey for collecting manufacturing energy consumption and related data from manufacturing establishments. The survey has the following characteristics:

- Quadrennial;
- Self-administered written questionnaire;
- Designed and sponsored by EIA and conducted by the U.S. Census Bureau;
- Mandatory by federal law;
- Confidential by Title 13 of the U.S Code;
- Statistical sample from list frame (15,500 sample cases in 2002).

The next MECS will be conducted in 2007 to collect data for 2006. As the MECS is quadrennial and relatively expensive, the importance of collecting useful, accurate, and comprehensive data becomes even more pronounced than with more frequent surveys.

The four-year gap requires that EIA re-examine the survey's format and content every time the survey is conducted. The condition of manufacturing changes as well as the technology available to them and to us, the survey designers and implementers.

For those reasons, and some that will be explained in the rest of this paper, we anticipate making a huge qualitative change in the way the MECS has been conducted in the past. The 2006 MECS will largely be an Internet-based electronic survey, with a paper form back-up for those that require it. The justification for this change is found in the performance statistics from the 2002 MECS.

The performance statistics data file has a wealth of "metadata" available in the Standard Economic Processing System (StEPS) database in which it is housed. For each data item

collected or derived, a flag indicates whether the item is reported, corrected by analyst intervention, or the result of using an alternate data source. The number of analyst corrections for collected data items gives an indication of data quality. Those questions that had particularly poor data quality using these criteria can be reworded or reconceptualized.

The 2002 MECS also had an electronic option for a certain class of respondents. Using Microsoft Excel, the written questionnaire was translated into a similar format that also had the capability of internally calculating results needed for completion in different parts of the questionnaire and would display warnings for data that failed price range checks. By comparing the metadata of particular items of this version with the written questionnaire data, we can determine whether we can expect better results using the electronic questionnaire in 2006.

Finally, the 2002 MECS tracked differential response rates by type of industry and size of establishments. In fact, the nonresponse follow-up was done in such a way as to maximize the reported data from those establishments that would have the greatest effect in minimizing the nonresponse adjustment of various cells. However, this approach favors spending resources to retrieve data from larger respondents at the expense of finding data from smaller ones. This may cause a nonresponse bias if the smaller respondents are qualitatively different from the larger ones. The 2006 MECS will examine adjusting the sample to favor larger establishments but increase the effort to retrieve smaller data from the smaller ones.

While doing the analysis in this paper and the initial preparations for the 2006, several questions arose for which we would be interested in the committee's input. They are:

1. Given that only 10 percent of the eligible respondents chose to use the electronic Excel form in 2002, and we have a self-selecting group, have we introduced bias into the comparisons? If so, are we still justified in reporting the results?
2. How do we boost participation in the Internet electronic form, given that respondents will have a back-up available to them?
3. How many onsite-edits should we have? Even if they're soft edits, do we fast exceed a common-sense limit?
4. What other performance statistics would be the most useful to track?
5. What are the risks associated with targeting nonresponse follow-up to larger establishments to minimize a nonresponse adjustment?
6. Should we adjust the sample even more to favor larger establishments and allow weights of smaller establishments to increase over the current maximum, given their higher rates of nonresponse?

Data Item Quality

Along with each item stored on the StEPS database for MECS, a *status flag* resides that indicates the origin of the entry. Flags found for 2002 are:

- R: Reported data; either through questionnaire or follow-up telephone or e-mail.
- A: Analyst Correction of Reporting Error; Obvious mistake correction, correction of units, correction of quantity based on expenditure, etc.

- I: Imputed data; missing data imputed by price or other information (rare)
- E: System edit; data changed due to systematic edit because of common mistake in reporting.
- S: Source data: An entirely different set of data was used instead of MECS for the item. In 2002, this usually meant the 2002 Economic Census—Manufacturing (ECM) for the chemical feedstocks, and the EIA-810, “Monthly Refinery Report” (Aggregated), for certain refinery data.

Figure 1 is a summary by percentage of status flags in for some items of interest in MECS. A quick scan of this and other similar charts is a fast way to make some judgments about both the quality of the data on the data base and the efficiency of asking the data on the questionnaire. A higher quality data item would normally have a higher percentage of “R” flags and fewer “A”, “I”, or “E” flags. The “S” flag is probably an indicator of good data quality but also indicates that the questionnaire item may need some improvement. The “R” flag may indeed be an indicator of good data quality, but it overstates the efficacy of our questions as both questionnaire reports and later follow-up contacts, in which a respondent gave actual data, are counted. However, StEPS does give a way to break this down further, although currently not without some extra manipulation. The data that are originally reported on the questionnaire are still resident along with the edited corrections that are used for final estimation and aggregation. By comparing the original version of the data with the edited, we can then further break out the “R” flags into two groups:

- The original data agrees with the edited indicating that no intervention was done; and
- The original data and the edited version do not agree which means the reported data came from a data editor/analyst follow-up.

Figure 1: Status Flags of Selected Data Items, Manufacturing Energy Consumption Survey, 2002

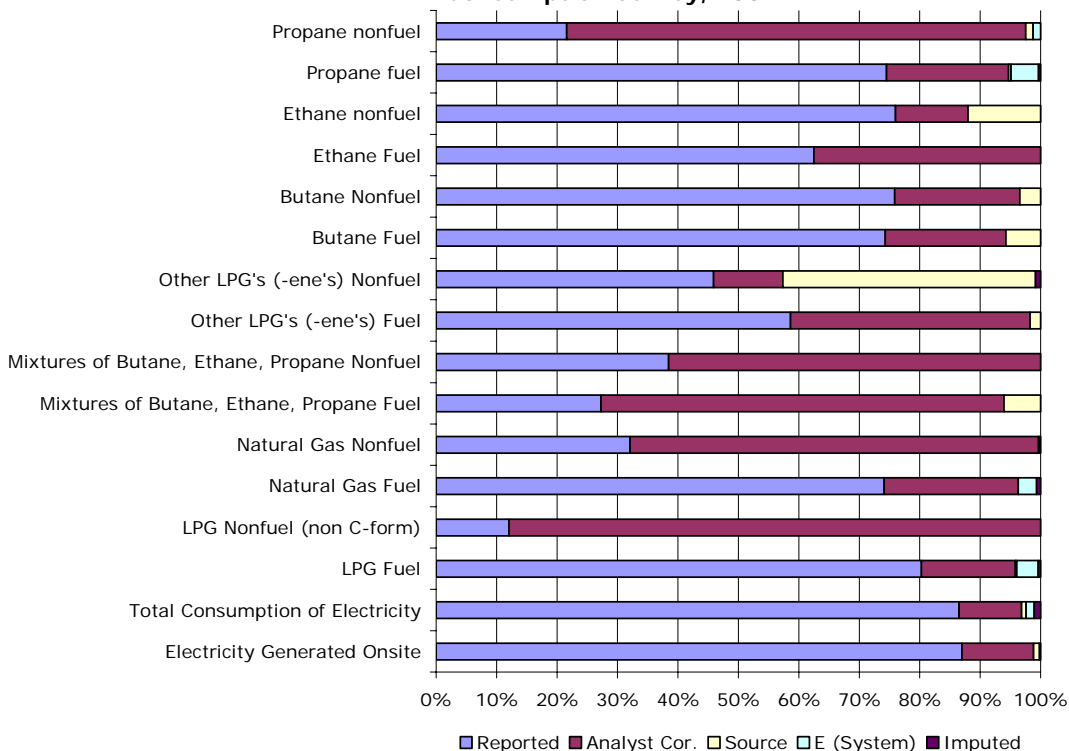


Table 1: Type of Status Flag as Percentage of Total, Manufacturing Energy Consumption Survey, 2002

	Reported	Analyst Correction	Source	E (System)	Imputed	Total Flags
Electricity Generated Onsite	87.0%	11.9%	1.0%	0.0%	0.2%	523
Total Consumption of Electricity	86.5%	10.3%	0.8%	1.3%	1.1%	11,199
LPG Fuel	80.3%	15.6%	0.2%	3.6%	0.4%	3,345
LPG Nonfuel (non C-form)	12.0%	88.0%	0.0%	0.0%	0.0%	183
Natural Gas Fuel	74.1%	22.2%	0.0%	3.0%	0.7%	9,057
Natural Gas Nonfuel	32.1%	67.6%	0.2%	0.0%	0.2%	623
Mixtures of Butane, Ethane, Propane Fuel	27.3%	66.7%	6.1%	0.0%	0.0%	33
Mixtures of Butane, Ethane, Propane Nonfuel	38.5%	61.5%	0.0%	0.0%	0.0%	13
Other LPG's (-ene's) Fuel	58.6%	39.7%	1.7%	0.0%	0.0%	58
Other LPG's (-ene's) Nonfuel	45.9%	11.5%	41.8%	0.0%	0.8%	122
Butane Fuel	74.3%	20.0%	5.7%	0.0%	0.0%	35
Butane Nonfuel	75.9%	20.7%	3.4%	0.0%	0.0%	29
Ethane Fuel	62.5%	37.5%	0.0%	0.0%	0.0%	8
Ethane nonfuel	76.0%	12.0%	12.0%	0.0%	0.0%	25
Propane fuel	74.5%	20.2%	0.4%	4.6%	0.3%	1,589
Propane nonfuel	21.6%	75.9%	1.2%	1.2%	0.0%	162

In fact, the proportion of R's that were the result of analyst interventions were usually below 10 percent of the total R's. One exception shown above was the LPG (Liquified Petroleum Gases) nonfuel from non C-form respondents.¹ Even the R's shown above are approximately 12 percent of the total for this data item. However, only 77 percent of those are non-intervention R's. What is truly noteworthy about this category was the number of analyst corrections required. Probably, these corrections represent conversions of reported data to zero as LPG nonfuel use outside of chemicals and petroleum refineries² would be quite rare.

Similarly, the status flags for the nonfuel use of propane represented a sizeable misreporting from industries that would normally not have such consumption. Often, respondents confuse nonfuel use with non-process consumption, especially when the non-process consumption is used in transportation equipment like forklifts. As a result, many of the analyst corrections for the nonfuel use of propane were recoding the nonfuel use into fuel consumption.

Natural gas used as a fuel is one of the most reported data items in the MECS. In 2002, 71 percent of the 200,000 establishments in manufacturing used natural gas as a fuel. The data for this item in Figure 1 represent approximately 9,000 out of the 15,500 sampled establishments. At that high level of reporting, it is important to minimize costly analyst intervention to any extent possible. Yet, as can be seen in Figure 1, 74 percent of the cases used in the final MECS dataset actually come from directly reported data (93 percent of those R's did not require analyst intervention.) The 22 percent "A" flags represent mostly conversions of the reporting units after failure of a price range check. Many of those edit changes could have been avoided if the respondents were warned of an edit failure while at the establishment site. Of course, the only way that could have occurred using a self-administered questionnaire would be if the instrument was electronic and could have the ability to compute price checks and warn respondents of failures.

A different type of problem emerges in the preponderance of "S" flags in the nonfuel use of different types of LPG. In Figure 1, the nonfuel use of "Other LPG's (-enes)," which includes propylene, ethylene, and butylene, has an especially high rate of that flag, 42 percent. The 2002 MECS and 2002 Economic Census—Manufacturing (ECM) covered the same data year for the first time in the history of the MECS. That identical coverage, both in time period and population, allowed for a comparison and exchange of data that was not available in previous MECS. The "S" flag for LPG nonfuel represents the substitution of ECM data for MECS when the MECS might have missing or questionable data. When data editors contacted the respondents to reconcile the discrepancy between the MECS and the ECM concerning the LPG nonfuel use, they discovered that Chemical respondents did not interpret the MECS reporting requirements correctly. They viewed the MECS as strictly an energy survey and that the use of LPG for nonfuel use was seen as out-of-scope, even though instructions for the MECS would suggest otherwise. Apparently, "nonfuel" or "feedstock" use was not seen as a "material input," the ECM

¹ The "C-Form" is sent to MECS respondents in energy-intensive industries with potentially complicated energy flows and nonstandard energy sources. All chemical plants would receive a C-form. In that version of the MECS, respondents are asked about specific types of LPG fuel or nonfuel use (e.g. butane and ethane). In Figure 1, the "LPG nonfuel (from non C-form)" would thus be from establishments outside the Chemical industry and so would be a rare and often incorrectly reported occurrence.

² The MECS does not collect nonfuel use of energy sources from petroleum refineries as it would duplicate data collection from other EIA data systems.

term that induced them to report correctly on that survey. Thus, as part of the 2006 MECS, we intend to use that term along side the other traditional MECS terms.

While status flags such as the ones used for the MECS are useful to show the quality of the data and point to potential problems in the questionnaire, they do not necessarily directly reflect failures in a data edit. Many of the analyst interventions were as a result of the failure of an edit that may be associated with a related or altogether different item. A well-trained analyst may then examine a MECS edit failure and, as a result, identify other errors that lead to the “A” status flag. However, the examination of the data status flags may ultimately lead to improved edit tests as well as changes in questionnaire mode and terminology.

The 2002 MECS Electronic Version of the C-Form

As an entrée into electronic data collection, EIA developed an Excel workbook version of the MECS C-Form³. The Excel version looked very much like the [written questionnaire](#) but it had added capabilities to:

- Check price ranges of major energy sources;
- Automatically calculate derived data items that normally the respondent would be required to do (e.g., total consumption of electricity);
- Automatically copy reported and derived data to later sections of the questionnaire when needed.

Although the Excel version of the C-Form had those beneficial capabilities, time and other constraints prevented it from optimal development. Not all desired onsite edits or screeners could be included. One major flaw was that the data from the Excel spreadsheet was not fully integrated with Census data capture routines. As such, after the respondent completed the questionnaire, he/she still had to print out or save the Excel output on a disk. The paper or disk had to be mailed to the same Census address to which the completers of the paper questionnaire were directed. Their responses would then be keyed the same as with the normal paper-and-pencil mode and thus be subject to the same potential for keying errors. The data from the electronic questionnaire for the 2006 MECS, on the other hand, will bypass data keying.

Using the most common data element for responders, total consumption of electricity⁴, the number of responses from the electronic version was 323, 8.9 percent of the total number of C-form responders to this item. However, the actual amount of total (first-use) consumption in Btu accounted for by the electronic responders was 12.4 percent.

Table 2 below compares the status flags of some of the common electricity variables for electronic and non-electronic reporters. It appears that respondents in these industries do fairly well in responding regardless whether they are using the electronic or paper versions. There was a price edit in place for the electronic responders but from the expenditure and purchased quantity electricity data, it appeared to have been somewhat extraneous. However, for the Total

³ The C-Form is sent to MECS respondents in the Wood Products, Paper, Chemicals, Petroleum and Coal Products (excluding petroleum refineries), and Iron and Steel industries. Those industries have complicated energy flows, are energy intensive, and/or are users of nonstandard energy sources.

⁴ This item must be present for a form to be considered a valid unit response.

Electricity Consumption variable, it is clear that there has been a noticeable effect for having an electronic form. Presumably, the feature that calculates the electricity consumption from the applicable reported data was a distinct advantage over having the respondents do it manually.

Table 2: Proportions of Status Flags for Selected Electricity Measures by Electronic and Paper responders.

Type of Reporter	Flag	Total Electricity Purchased	Total Electricity Expenditures	Total Electricity Consumption
Electronic	R	96.0%	97.5%	96.3%
	A	3.4%	1.9%	3.1%
	Other	0.6%	0.6%	0.6%
Paper	R	92.5%	96.2%	85.7%
	A	6.2%	2.8%	11.9%
	Other	1.3%	1.0%	2.3%

Unlike electricity which is normally recorded in some multiple of kilowatthours, natural gas is billed and recorded in many different units, including thousand cubic feet (MCF), hundreds of cubic feet (CCF), therms, or million Btu. This causes an added problem for MECS respondents who must convert their natural gas records into MCF, the unit required by MECS. Although the MECS does provide conversion factors and formulas to do this, many respondents have trouble doing this correctly.

An onsite price check would presumably be very helpful in catching many of those reporting errors. Table 3 shows the percentage of status flags for three of the major natural gas variables. As in the electricity table above, the status flags are shown separately for electronic and paper form reporters. For total natural gas expenditures, there is no real difference in using the electronic or paper form. In fact, analysts tend to use the reported expenditures to guide editing of related data. Yet, for purchased quantities there appears to be a discernable difference in the need for analyst or other intervention.⁵ The inference could be made that the onsite price check did help with the quality of the reporting. However, we have no way of definitively determining whether this was true since it is possible that the electronic respondents also happened to be a better set of reporters.

⁵ Using the methodology described previously comparing reported and edited versions of the data, the percentage of R's that actually were a result for all three of these variables was over 93 percent.

Table 3: Proportions of Status Flags for Selected Natural Gas Measures by Electronic and Paper responders.

Type of Reporter	Flag	Total Natural Gas Purchased	Total Natural Gas Expenditures	Total Natural Gas Fuel Consumption
Electronic	R	87.8%	96.7%	88.6%
	A	11.3%	3.3%	8.1%
	Other	0.9%	0.0%	3.3%
Paper	R	76.0%	95.9%	74.9%
	A	21.7%	3.6%	16.8%
	Other	2.3%	0.6%	8.3%

In the natural gas fuel consumption variable, not only might a price edit warning have an effect, but so might also the automatic copying of the natural gas consumption data from one section to another. The variable shown in Table 3 is actually the total consumption used as the basis for determining end-uses. In fact, the 8.3 percent shown as “Other” for paper responders was actually an “E” flag, indicating a systematic edit without human analyst intervention, which was the most efficient way of performing those transfers after data is received. As shown in Table 3, the “Other” flag preponderance was much greater for paper-form responders.

The comparison of many of the other purchase and consumption variables tends to become problematic as the numbers of establishments for whom they apply can become very small. However, it did appear that electronic reporting was at least as good or more favorable in all cases examined.

Future Electronic Reporting

As mentioned previously, the 2002 Excel version of the C-form did not have all the features that we would have introduced had there been more time for development. First, as already discussed, the Excel version was not integrated with StEPS so that input had to be done using the same keying procedures used for paper forms. Work has already started for the 2006 MECS electronic questionnaire using Census Taker to ensure that this integration will take place.

Other intended features of the 2006 MECS electronic questionnaire:

- All respondents will be eligible to use it. Indeed, they will be highly encouraged to do so and paper will be made available only as a back-up;
- Industry classification (i.e., NAICS code) will be used as a screener to help build the form on the fly. That screening will allow the respondent to focus in only on the energy sources most relevant to that type of industry.
- Check boxes or other means will be used to give the respondent an opportunity to report in his or hers most convenient units. Edit price checks will be developed for each of the unit choices.
- The questionnaire will be developed and formatted with screen input in mind, using the best available practices.

- As the electronic version will be customized to the respondent, we hope respondents in smaller establishments will be more likely to respond, as they will not be put off by the apparent size of the written version. Essentially, the many sections of the written questionnaire that would not apply, would never be seen by the respondent.

Differential Response Rates

As in the case of other establishment surveys, the problem of nonresponse is the MECS has been growing. Greater emphasis on labor productivity and other efficiencies to maintain a competitive edge has led manufacturers to reduce or shut down non-core functions such as would be required to complete surveys such as the MECS. Even laws that require compliance, such as the one in place for conducting the MECS, do not have the influence they once had. Perhaps this is a reflection of attitudes toward government and surveys in general or a realization that mandatory compliance was very unlikely to be enforced.

On the other side, the budgets necessary to conduct large government surveys are also being squeezed. In the case of the MECS, we have had to reduce the sample size in each of the last two cycles of the MECS and hold the line on the number of questions. There just is not extra money to use for extended nonresponse follow-up. In order to best use the resources available to us, the non-response follow-up for the 2002 MECS was targeted to have the greatest effect on the survey data quality.

The MECS is a statistical sample with establishment inclusion probability proportional to an energy measure of size (MOS) within an industry cell.⁶ The nonresponse adjustment for 2002 is done separately for certainty establishments (with inclusion probabilities of 1) and noncertainty establishments. In adjustment cells for certainty establishments, those establishments that have higher MOS also have greater influence on the final nonresponse adjustment.⁷ Further, cells having higher overall MOS also have greater influence on the final MECS energy estimates. For those reasons, it made sense to target nonresponse follow-up to those large establishments in cells that have the greatest influence on our estimates.

Table 4 shows the results of that target nonresponse follow-up. Coverage rate is the proportion of total MOS for any cell accounted for by the MOS of the respondents. A high coverage rate will minimize the size and effect of the nonresponse adjustment to the sampling weights. That will in turn reduce the size of the variance of the estimates.

⁶ The situation is complicated by needing to alter some probabilities of selection for the likely presence of feedstock or other characteristics not captured by the energy MOS.

⁷ For more details on the MECS sampling frame, sample design, estimation procedures, and nonresponse adjustments see, [2002 Manufacturing Energy Consumption Survey Methodology and Data Quality: Survey Design, Implementation, and Estimates](#) on the EIA Website.

Table 4: Response and Coverage Rates by Cell Size and Type.

Cell Size	Target Coverage Rate	Type	Number in Cell	Reporters in Cell	Non-reporters in Cell	Response Rate	Achieved Average Coverage Rate
Small	70%	Certainty	73	50	23	68%	85%
		Noncertainty	134	89	45	66%	
Medium-Small	75%	Certainty	566	457	109	81%	84%
		Noncertainty	856	600	256	70%	
Medium	80%	Certainty	1,751	1,452	299	83%	87%
		Noncertainty	1,987	1,457	530	73%	
Medium-Large	85%	Certainty	1,884	1,624	260	86%	87%
		Noncertainty	4,511	3,294	1,219	73%	
Large	90%	Certainty	552	483	69	88%	92%
		Noncertainty	3,212	2,422	790	75%	
Grand Total			15,526	11,928	3,600	77%	90%

As can be seen in Table 4, the achieved coverage rate exceeded the target in all cases. This was true, even though the response rate in terms of valid submissions was relatively low.

For the 2006 MECS, EIA and Census intend to use the same targeted approach to nonresponse as was used in the 2002 MECS. We are also exploring the possibility of sampling even at a lower rate from smaller establishments, given that their response rates are generally lower.

Conclusions

It is important to maintain or even improve the overall quality of the MECS, given the 4-year gap between cycles and the relevance of the data themselves increases in an economy that increasingly can not be adequately described by supplier data alone. As a start, the 2006 MECS will shift to an Internet-based electronic reporting method for its main mode of collection. To achieve a high use rate of the electronic questionnaire, we intend to draw the sample earlier in the cycle. This will enable us to have an early pre-mailing so that we can identify the individuals who will be responsible for completing the survey. Much of the 2002 MECS late- and non-response was due to not having the right contact persons at the establishment in the beginning. As described above, we may adjust the sample away from the smaller establishments to account for their historic smaller response rates.

After the data are received, we intend to take further advantage of the available metadata and performance statistics. In fact, the status flags reported on in the paper will be amended to further define the occurrence and type of analyst intervention.

Questions for the Committee

1. Given that only 10 percent of the eligible respondents chose to use the electronic Excel form in 2002, and we have a self-selecting group, have we introduced bias into the comparisons? If so, are we still justified in reporting the results?
2. How do we boost participation in the Internet electronic form, given that respondents will have a back-up available to them?
3. How many onsite-edits should we have? Even if they're soft edits, do we fast exceed a common-sense limit?
4. What other performance statistics would be the most useful to track?
5. What are the risks associated with targeting nonresponse follow-up to larger establishments to minimize a nonresponse adjustment?
6. Should we adjust the sample even more to favor larger establishments and weights of smaller establishments increase over the current maximum, given their higher rates of nonresponse?